

北京丁香质体基因组研究

鲍海娟^{1,2},徐春柳¹,王燕燕¹,谷小龙¹,刘晶^{1,2}

(1.青海民族大学 a.生态环境与资源学院;b.青海省特色经济植物高值化利用重点实验室;
c.药用植物资源学凯瑞研究生工作站,西宁 810007;

2.青海奔盛草业有限公司 农业农村部西北耕地保育及边际土地改良重点实验室,青海 德令哈 817000)

摘要:为了解北京丁香的质体基因组特征,本研究分析了北京丁香的质体基因组结构和基因构成,并联合其它 35 个木犀科物种和 2 个泡桐科物种进行了系统发育分析。结果表明:1) 北京丁香的质体基因组共编码 131 个基因,包含 36 个 tRNA 基因、8 个 rRNA 基因和 87 个 mRNA 基因,分别占有注释基因的 27.48%、6.11%和 66.41%;2) 在该质体基因组的密码子组成中,RSCU 值大于 1 的密码子有 30 个,31 个密码子的 RSCU 值小于 1,且密码子有以 G 和 C 结尾的使用偏好性;3) 北京丁香质体基因组中共识别出 178 个正向重复序列和 177 个回文重复序列;4) NJ 系统发育树表明北京丁香与丁香属植物暴马丁香亲缘关系较为密切,丁香属植物与女贞属植物亲缘关系密切。

关键词:北京丁香;木犀科;质体;系统发育树

中图分类号:S685.26 **文献标志码:**A **文章编号:**1673-8020(2024)04-0360-10

北京丁香(*Syringa reticulata* subsp. *pekinensis*) 是木犀科(Oleaceae)植物网脉丁香(*Syringa reticulata*)的亚种^[1]。为落叶小乔木或大灌木,部分植株可高达 10 m,树皮呈灰棕色或者褐色、细长、向外伸展,有明显的皮孔,新生枝条有柔毛。叶片形状是常见的椭圆形或卵形,花冠呈白色,其长度约为 3~4 mm,呈辐射状^[1-3]。北京丁香花期为 5—8 月,因花香怡人,常被用于提炼精油、制作空气清新剂和化妆品等^[4]。其挥发油具有抗菌、消炎、抑制肿瘤、镇痛以及止泻等多种功效^[1,5-6],其嫩叶亦可作茶饮,有清热解毒功效^[7]。北京丁香具有良好的耐寒性、耐干旱性、耐高温性等抗逆性,因此被广泛应用于城乡庭院绿化中^[8-9]。

质体是绿色植物所特有的半自主性细胞器,包含白色体、有色体和叶绿体 3 种类型,三者共享一套遗传系统,即质体基因组^[10-11]。根据现有 NCBI 公开数据显示,植物质体基因组大小在 3.84~1320.61 kb 之间(<https://www.ncbi.nlm.nih.gov/nuccore>),其中高湖泊红球藻(*Haematococcus lacustris*)的质体基因组最大(1320.61 kb),一个拟胶毛藻相似种(*Chaetophoropsis* cf.

attenuata)的质体基因组最小(3.84 kb)。而在陆生植物中,质体基因组大小一般在 120~180 kb 之间,大约编码 60~100 个蛋白编码基因(Protein coding gene, CDS)、25~32 个转运 RNA(Transfer RNA, tRNA)及 4 个核糖体 RNA(Ribosomal RNA, rRNA)^[12],拥有丰富的遗传信息,容易获取,且测序难度较小^[13]。多数植物的质体 DNA 为双链环状,但伞藻(*Acetabularia*)等少数植物为线状^[14]。目前公认的质体基因组有母系遗传、父系遗传和双亲遗传 3 种遗传方式^[15],被子植物的质体基因组一般遵循母系遗传,但也有少部分遵循父系遗传,如胡萝卜属^[16]、番薯属^[17]和甘草属^[18]等,裸子植物一般遵循父系遗传^[19]。总体来说,质体基因组具有低碱基替换率、母系遗传、基因组序列存在保守性、分子量小等优点,使得其在植物的系统进化、物种分类与鉴定等研究中得到广泛的应用^[20]。吴民富等^[21]采用高通量测序技术开展猫尾草叶绿体基因组测序,研究发现猫尾草与同属狸尾豆的亲缘关系最近^[22]。尚明越等^[22]通过采用 Illumina 技术对紫皮石斛进行测序,发现紫皮石斛叶绿体基因组呈典型四分体结

收稿日期:2024-03-09;修回日期:2024-06-06

基金项目:青海省重大科技专项(2023-NK-A3)

通信作者简介:刘晶(1983—),女,副教授,硕士研究生导师,博士,研究方向为植物种质开发与利用。E-mail:654138133@qq.com

构,并且与兜唇石斛亲缘关系最近。

本研究拟对北京丁香的质体基因组进行测序、组装和注释等生物信息学分析,并根据注释好的完整质体基因组绘制图谱,随后分析密码子偏好性、散在重复序列以及简单重复序列,最后结合已公开的35种木犀科和2种泡桐科植物的质体基因组序列构建系统发育树。基于此,揭示丁香属植物亲缘关系,探明丁香属植物起源、进化及分类阶元的系统发育,为今后北京丁香的分子育种提供遗传学背景信息,也为北京丁香在园林绿化中的应用提供理论依据,更为北京丁香未来的进化和生态学研究提供有效的支撑。

1 实验材料与方法

1.1 材料来源

在青海省西宁市植物园(36°37'N, 101°46'E, 海拔2250 m)采集北京丁香成熟且无病虫害的叶片放入茶包中,加入硅胶干燥剂,带回实验室后于-20℃保存。另取部分带有健康叶片的北京丁香枝条制成标本,保存于中国科学院西北高原生物研究所青藏高原生物标本馆中。

1.2 DNA的提取、测序与组装

用2×CTAB法提取北京丁香总的DNA,将提取到的符合相关要求的DNA基于Illumina NovaSeq平台进行高通量测序,此项工作交由南京集思慧远生物科技有限公司完成。为保证后续实验结果的科学性,测序得到的原始数据需进行质量控制,截除raw reads中带有测序接头的序列、引物序列以及质量值小于5的碱基数占整个read的50%以上的序列,从而得到最终待分析的clean reads。随后将其公司自建的质体基因组数据库用bowtie2 v2.2.4(<http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>)进行比对,随后将获得的cpDNA测序序列利用fastp软件对数据进行质量控制^[23]。最后采用SPAdes软件将质体基因组进行组装。

1.3 质体基因组注释、基因组图谱绘制

将测序所获得的北京丁香质体基因组序列(fasta格式),应用在线注释软件GeSeq进行基因

注释^[24]。具体步骤为:上传fasta格式的cpDNA序列,参数选择Circular、Sequence source;勾选Chloë v0.1.0,选择CDS、tRNA、rRNA,剩余参数默认;点击submit获得北京丁香质体基因组的注释结果,人工校对后保存。最后利用在线软件OGDRAW(<https://chlorobox.mpimp-golm.mpg.de/OGDraw.html>)进行质体基因组图谱的绘制^[25]。

1.4 同义密码子相对使用度(relative synonymous codon usage, RSCU)分析

密码子在遗传信息翻译成蛋白质的过程中起着至关重要的作用,氨基酸(甲硫氨酸和色氨酸除外)至少带有两个以上的密码子^[26]。一方面密码子具有简并性,因为每个氨基酸对应多种密码子,所以有着多个密码子编码相同氨基酸的现象^[27]。另一方面密码子还具有使用偏好性,是指同义密码子不仅在不同基因组中的使用频率不同,而且在同一基因组的不同基因中也不相同^[28]。而密码子的这种偏好性被认为是物种进化的特征之一,是遗传漂变、物种突变、自然选择、tRNA的丰度等综合作用的结果^[29]。研究密码子偏好性不仅对于探究生物的进化发展有帮助,而且对于外源基因的表达有十分重要的意义^[30-31]。本研究中密码子的运用频率和RSCU值利用CodonW软件进行测试^[32]。

1.5 重复序列分析

散在重复序列和串联重复序列是质体基因组重复序列的两种表达形式。基因散在重复序列指基因的重复单位散布在整个基因组中的序列,其类型有4种,即反向重复(R型)、正向重复(F型)、回文重复(P型)和完全重复(C型)^[33-34]。散在重复序列的分析在分子标记、动植物育种等方面的研究中起着非常重要的作用^[35]。本文使用在线软件REPuter鉴定北京丁香质体基因组的散在重复序列^[36-37]。

串联序列包括微卫星DNA、小卫星DNA以及卫星DNA。微卫星序列(microsatellite DNA)又称为简单重复序列(simple sequence repeats, SSR),是由小于10 bp的DNA组成的串联重复序列^[36]。基于叶绿体基因组开发的标记称为叶绿体微卫星(chloroplast simple sequence repeatscp, cpSSR),cpSSR的核心重复序列类型主要包括单核苷酸重复、二核苷酸重复、三核苷酸重复和四核

苷酸重复^[38]。cpSSR 技术不仅继承了 SSR 分子标记重复性好、多态性高等诸多优势,而且它还兼具叶绿体基因组高度保守、单亲遗传特性、独立进化和对 DNA 样本质量的较低要求等优点,因此这一技术已被广泛应用于多个植物研究领域^[39]。本研究中 cpSSR 分析用 MISA 在线软件(<http://webblast.ipk-gatersleben.de/misa/>)完成,SSR search parameters 设置为 1-8、2-5、3-3、4-3、5-3 和 6-3,其余参数设置默认值^[40]。

1.6 系统演化分析

从 NCBI 数据库中选取 35 种木犀科植物和 2 种泡桐属植物(外类群)的质体基因组序列进行系统发育分析。用于构建系统发育树的木犀科植物包括:5 种女贞属(*Ligustrum*)植物、14 种丁香属(*Syringa*)植物、3 种流苏树属(*Chionanthus*)植物、2 种岛蜡树属(*Nestegis*)植物、木樨榄属(*Olea*)植物、鳞瓣榄属(*Haenianthus*)植物、泽蜡树属(*Forestiera*)植物、总序桂属(*Phillyrea*)植物、大苞榄属(*Picconia*)植物、桫欏属(*Priogymnanthus*)植物、爪瓣榄属(*Hesperelaea*)植物、木樨属(*Osmanthus*)植物、元春花属(*Schrebera*)植物、桫欏属(*Fraxinus*)、翘果连翘属(*Abeliophyllum*)植物各 1 种,外类群泡桐科植物包括 2 种泡桐属(*Paulownia*)植物(表 1)。基因组序列对齐采用 MAFFT 的在线版本(online version of addfragments)完成。系统进化树的构建采用 MAGAX 软件的邻接法(Neighbor-Joining, NJ)完成。

2 结果与分析

2.1 质体基因组基本特征

2.1.1 测序及组装结果

利用 Illumina NovaSeq 平台对北京丁香的总 DNA 进行测序,测序度长为 PE 150。测序结果获得 ReadSum 的数量为 33 758 871, Clean Data 的总碱基数约为 9.43 Gb, GC 含量为 34.85%, Q20 的含量为 97.27%, Q30 的含量为 92.49%。

2.1.2 北京丁香质体基因组基本特征

经过测序、质控、组装得到全长为 155 819 bp

的北京丁香质体基因组(NCBI 登录号:MN901632.1)的完整序列。经注释,发现该质体基因组结构与许多其他被子植物结构相似,均为四分的环状结构,分别为大单拷贝区域(large single copy, LSC)、小单拷贝区域(small single copy, SSC)、反向重复区(inverted repeat region, IR),其中 IRA 和 IRB 为反向重复序列的组成部分,LSC、SSC、IRA、IRB 长度分别为 87 055、40 064、14 350 和 14 350 bp。两个 IR 区域由 LSC 和 SSC 分隔开。质体基因组序列的 GC 含量为 38.06%。北京丁香的质体基因组注释了 36 个 tRNA 基因,8 个 rRNA 基因和 87 个蛋白质编码基因,共 131 个。三者分别占有注释基因的 27.48%、6.11% 和 66.41% (表 2,图 1)。

2.1.3 RSCU 结果

北京丁香质体基因组密码子使用率结果表明(表 3,图 2):有 30 个密码子的 RSCU 值大于 1,其中有 28 个密码子以 A/U 结尾;31 个密码子的 RSCU 值小于 1,其中有 29 个密码子以 G/C 结尾;且 Trp(色氨酸)无密码子偏好性,因为它的 RSCU 值等于 1。

2.1.4 散在重复序列分布

使用 REPuter 软件对北京丁香质体基因组进行散在重复序列调查(图 3)。在北京丁香质体基因组中仅发现正向重复序列(F 型)和回文重复序列(P 型),无反向重复序列(R 型)和完全重复(C 型)序列。其中 F 型序列 178 个,P 型 177 个。散在重复序列长度大多集中在 15~22 bp 之间,占比约为 85.1%。无论是 P 型还是 F 型,都很少有长度大于 22 bp 的重复序列,而长度超过 1088 bp 的重复序列仅有 2 个,为回文重复序列。

使用 MISA 软件进行 cpSSR 的分析(图 4)。cpSSR 在整个基因组中并非平均分布,其分布情况如下:LSC 区域中有 142 个(65.10%),其中包括在 Exon 中 34 个、Intron 中 25 个,以及 Intergenic 中 83 个;SSC 区域中有 50 个(22.90%),其中包括在 Exon 中 23 个、Intron 中 5 个,以及 Intergenic 中 22 个;IR 区域中有 26 个(11.90%),其中包括在 Exon 中 18 个、Intron 中 2 个,以及 Intergenic 中 6 个。

表1 参与系统发育树构建的物种信息
Tab.1 Species information involved in the construction of phylogenetic trees

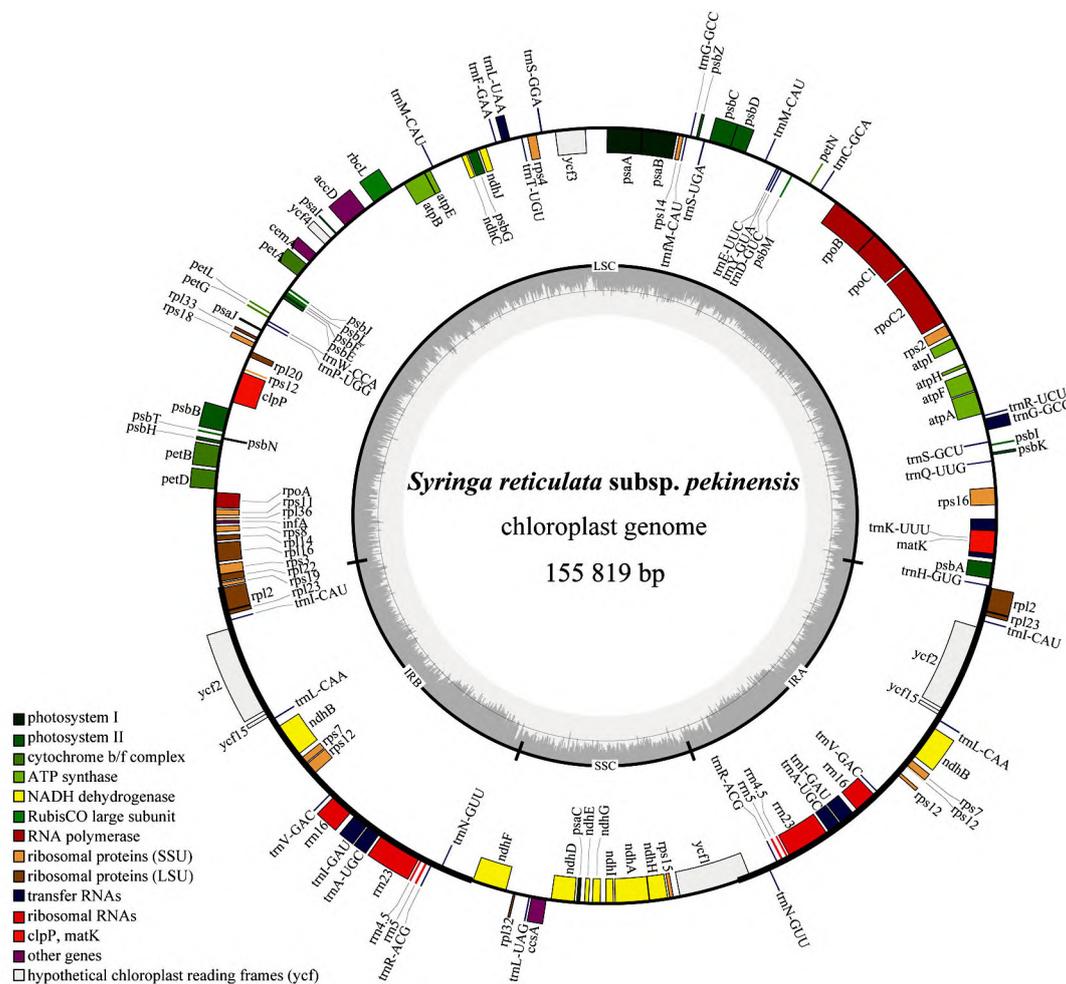
科名	属名	物种名称	拉丁名	登录号
木犀科	女贞属	细女贞	<i>Ligustrum gracile</i>	MH817914.1
		女贞	<i>Ligustrum lucidum</i>	MH394207.1
		日本女贞	<i>Ligustrum japonicum</i>	MH817915.1
		小叶女贞	<i>Ligustrum quihoui</i>	MN723863.1
		欧洲女贞	<i>Ligustrum vulgare</i>	MH817874.1
	丁香属	暴马丁香	<i>Syringa reticulata</i> subsp. <i>amurensis</i>	MW525283.1
		北京丁香	<i>Syringa reticulata</i> subsp. <i>pekinensis</i>	MN901632.1
		小叶巧玲花	<i>Syringa pubescens</i> subsp. <i>microphylla</i>	MH817881.1
		小叶巧玲花	<i>Syringa pubescens</i> subsp. <i>microphylla</i>	MT872641.1
		垂丝丁香	<i>Syringa komarowii</i> subsp. <i>reflexa</i>	MT648823.1
		红丁香	<i>Syringa villosa</i>	OL414766.1
		春川丁香	<i>Syringa fauriei</i>	OR209257.1
		羽叶丁香	<i>Syringa pinnatifolia</i>	OM912776.1
		辽东丁香	<i>Syringa wolfii</i>	MN901631.1
		云南丁香	<i>Syringa yunnanensis</i>	MH817943.1
		波斯丁香	<i>Syringa persica</i>	MH817880.1
		羽叶丁香	<i>Syringa pinnatifolia</i>	MG917095.1
		紫丁香	<i>Syringa oblata</i>	MT025818.1
	欧丁香	<i>Syringa vulgaris</i>	MG255768.1	
	流苏树属	短序流苏树	<i>Chionanthus brachystachys</i>	MK299388.1
		腋花流苏树	<i>Chionanthus axillari</i>	MH817882.1
		新几内亚流苏树	<i>Chionanthus brassii</i>	MH817883.1
	岛腊树属	岛腊树	<i>Nestegis apetala</i>	MG255758.1
		黑岛腊树	<i>Nestegis cunninghamii</i>	MH817916.1
	翅果连翘属	翅果连翘	<i>Abeliophyllum distichum</i>	KT274029.1
	元春花属	元春花	<i>Schrebera alata</i>	MH817939.1
	栲属	美国白栲	<i>Fraxinus americana</i>	MH817907.1
	木樨榄属	木樨榄	<i>Olea europaea</i>	GU931818.1
鳞瓣榄属	鳞瓣榄	<i>Haenianthus salicifolius</i>	MH817913.1	
总序桂属	总序桂	<i>Phillyrea latifolia</i>	MH817936.1	
大苞榄属	亚速尔大苞榄	<i>Picconia azorica</i>	MH817937.1	
木樨属	丹桂	<i>Osmanthus austrocaledonicus</i>	MK299397.1	
栲榄属	栲榄	<i>Priogymnanthus hasslerianus</i>	MH817879.1	
泽蜡树属	佛州泽蜡树	<i>Forestiera segregata</i>	MH817906.1	
爪瓣榄属	爪瓣榄	<i>Hesperelaea palmeri</i>	LN515489.1	
泡桐科	泡桐属	毛泡桐	<i>Paulownia tomentosa</i>	KP718624.1
	泡桐属	毛泡桐	<i>Paulownia tomentosa</i>	MK875778.1

表2 北京丁香质体基因组注释基因信息
Tab.2 Annotated gene information of *Syringa reticulata* subsp. *pekinensis* plastid genome

基因功能分类	基因种类	基因名称	基因数量/个
自我复制相关基因	核糖体 RNA 基因	<i>rrn16_{x2}, rrn23_{x2}, rrn4.5_{x2}, rrn5_{x2}</i>	8
	转运 RNA 基因	<i>trnH-GUG, trnK-UU, trnQ-UUG, trnS-GCU, trnG-GCC_{x2}, trnR-UCU, trnC-GCA, trnD-GUC, trnY-GUA, trnE-UUC, trnM-CAU_{x2}, trnS-UGA, trnM-CAU, trnS-GGA, trnT-UGU, trnL-UAA, trnF-GAA, trnW-CCA, trnP-UGG, trnL-CAA_{x2}, trnV-GAC_{x2}, trnI-GAU_{x4}, trnA-UGC_{x2}, trnR-ACG_{x2}, trnN-GUU_{x2}, trnL-UAG</i>	36
	RNA 聚合酶亚基基因	<i>rpoA, rpoB, rpoC1, rpoC2</i>	4
	核糖体小亚基	<i>rps16, rps2, rps14, rps4, rps18, rps12_{x2}, rps11, rps8, rps3, rps19, rps7_{x2}, rps15</i>	14
	核糖体大亚基	<i>Rpl33, rpl20, rpl36, rpl14, rpl16, rpl22, rpl12, rpl23_{x2}, rpl32, rpl2</i>	11

表 2 (续)
Tab.2 (Continued)

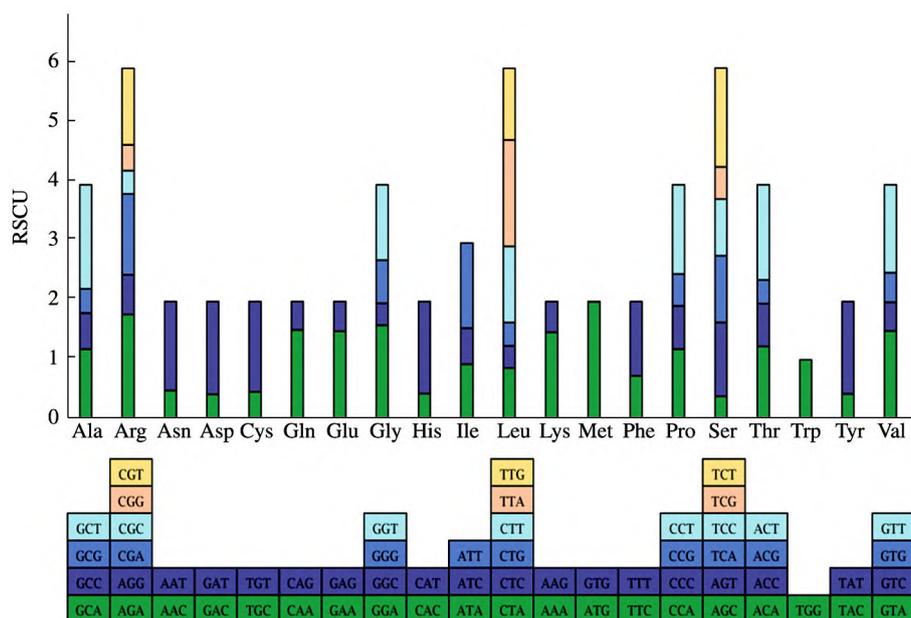
基因功能分类	基因种类	基因名称	基因数量/个
光合作用相关基因	光合系统 I 基因	<i>psaB, psaA, psaI, psaJ, psaC</i>	5
	光合系统 II 基因	<i>psbA, psbK, psbI, psbM, psbD, psbC, psbZ, psbG, psbJ, psbL, psbF, psbE, psbB, psbT, psbN, psbH</i>	16
	RubisCO 大亚基	<i>rbcL</i>	1
	细胞色素 b/f 复合物	<i>petN, petA, petL, petG, petB, petD</i>	6
	ATP 合成酶	<i>atpA, atpF, atpH, atpI, atpE, atpB</i>	6
	NADH 脱氢酶	<i>ndhJ, ndhC, ndhB, ndhF, ndhD, ndhE, ndhG, ndhI, ndhA, ndhH, ndhB</i>	11
其他基因	蛋白酶	<i>clpP</i>	1
	成熟酶	<i>matK</i>	1
	囊膜蛋白基因	<i>cemA</i>	1
	乙酰辅酶 A 羧化酶亚基基因	<i>accD</i>	1
	转录起始因子	<i>infA</i>	1
	C 型细胞色素合成基因	<i>ccsA</i>	1
未知功能基因	叶绿体阅读框	<i>Ycf3, ycf4, ycf2_{x2}, ycf5_{x2}, ycf1</i>	7



注:圈内侧和外侧分别代表反向编码和正向编码的基因;灰色圈代表 GC 含量。LSC:长单拷贝区;SSC:短单拷贝区;IRA, IRB:方向互补区。

图 1 北京丁香质体基因组图谱

Fig.1 Plastid genome map of *S. reticulata*



注:柱状图中柱子的高度表示该氨基酸全部同义密码子 RSCU 值的总数,方块表示每种氨基酸的全部同义密码子。

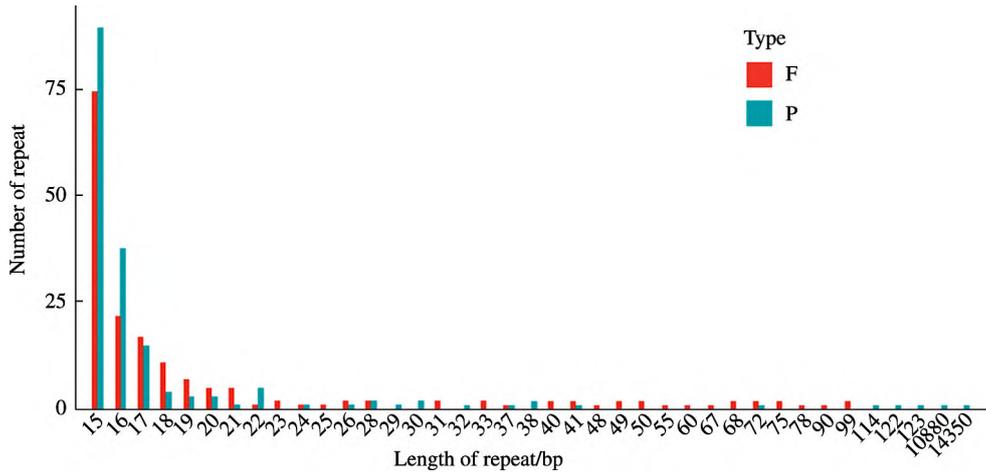
图2 RSCU 柱状图

Fig.2 RSCU histogram

表3 北京丁香质体基因组密码子使用率

Tab.3 The relative synonymous codon usage of the *Syringa reticulata* subsp. *pekinensis* plastid genome

氨基酸	密码子	序号	RSCU	氨基酸	密码子	序号	RSCU
ALA	GCA	406	1.179 2	ASN	AAC	279	0.468 6
ALA	GCC	214	0.621 6	ASN	AAU	912	1.531 4
ALA	GCG	143	0.415 2	PRO	CCA	323	1.176 8
ALA	GCU	614	1.783 6	PRO	CCC	204	0.743 2
CYS	UGC	63	0.448 4	PRO	CCG	152	0.553 6
CYS	UGU	218	1.551 6	PRO	CCU	419	1.526 4
ASP	GAC	207	0.404 6	GLIN	CAA	671	1.511 2
ASP	GAU	816	1.595 4	GLIN	CAG	217	0.488 8
GLU	GAA	995	1.487 2	ARG	AGA	457	1.778 4
GLU	GAG	343	0.512 8	ARG	AGG	174	0.676 8
PHE	UUC	511	0.723 2	ARG	CGA	357	1.389 0
PHE	UUU	902	1.276 8	ARG	CGC	102	0.396 6
GLY	GGA	697	1.589 6	ARG	CGG	114	0.443 4
GLY	GGC	165	0.376 4	ARG	CGU	338	1.315 2
GLY	GGG	326	0.743 6	SER	AGC	121	0.368 4
GLY	GGU	566	1.290 8	SER	AGU	417	1.269 0
HIS	CAC	123	0.418 4	SER	UCA	376	1.144 2
HIS	CAU	465	1.581 6	SER	UCC	319	0.970 8
ILE	AUA	647	0.920 4	SER	UCG	182	0.553 8
ILE	AUC	438	0.623 1	SER	UCU	557	1.695 0
ILE	AUU	1 024	1.456 5	THR	ACA	397	1.224 4
LYS	AAA	946	1.466 6	THR	ACC	239	0.737 2
LYS	AAG	344	0.533 4	THR	ACG	132	0.407 2
LEU	CUA	376	0.858 0	THR	ACU	529	1.631 6
LEU	CUC	165	0.376 2	VAL	GUA	507	1.494 4
LEU	CUG	176	0.401 4	VAL	GUC	165	0.486 4
LEU	CUU	573	1.307 4	VAL	GUG	173	0.510 0
LEU	UUA	800	1.825 2	VAL	GUU	512	1.509 2
LEU	UUG	540	1.231 8	TRP	UGG	455	1.000 0
MET	AUG	590	1.993 2	TYR	UAC	188	0.410 4
MET	GUG	2	0.006 8	TYR	UAU	728	1.589 6



注:横坐标代表散在重复序列的类型,纵坐标代表散在重复序列的数量。P 表示回文重复,D 表示正向重复。

图 3 北京丁香质体基因组散在重复序列分析

Fig.3 Analysis of Interspered repeated sequence in plastid genome of *S. reticulata* subsp. *pekinensis*

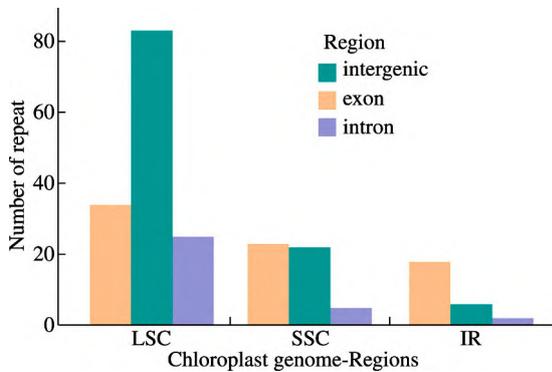


图 4 北京丁香 cpSSR 分析

Fig.4 cpSSR analysis of *S. reticulata* subsp. *pekinensis*

2.2 系统发育分析

通过构建木犀科植物和泡桐科植物的系统发育树,获取北京丁香的亲缘关系,将其和 35 种木犀科和 2 种泡桐科植物的质体基因组采用 NJ 法构建系统进化树,进行亲缘性关系分析。系统发育分析结果显示:绝大多数分支都获得了 100%的支持率,且同属植物聚为一支,进一步表明研究结果的准确性和可靠性。系统发育树可以分为 2 个大支。第一大支是所有的木犀科植物聚为一支,其中丁香属植物、女贞属、流苏树属和岛腊树属各聚为一支,所有的丁香属植物和女贞属植物又聚为一支,这表明与丁香属植物亲缘关系最近的是女贞属植物;北京丁香与暴马丁香(*S. reticulata* subsp. *amurensis*)亲缘关系最为密切,NJ 进化树表明二者以 100%的强支持率聚为一小支(图 5)。第二大支为外类

群泡桐科植物聚为一支(图 5)。

3 讨论

本研究对北京丁香质体基因组进行了分析,并基于质体基因组构建了系统进化树。

研究表明,北京丁香质体基因组的序列全长为 155 819 bp,其中 LSC、SSC、IRA、IRB 长度分别为 87 055、40 064、14 350 和 14 350 bp,其结构与许多其他被子植物相同均为四分的结构,这与前人研究的暴马丁香^[41]、辽东丁香^[42]、四川丁香^[43]、紫丁香^[44]的结论一致。北京丁香质体基因组的 GC 含量为 38.06%,这一结果和木樨科近缘物种相近^[45]。通过对北京丁香 20 种氨基酸密码子偏好性使用研究发现,有 30 个密码子 RSCU 值大于 1,其中 28 个密码子以碱基 A/U 结尾;31 个密码子的 RSCU 值小于 1,其中有 29 个密码子以碱基 G/C 结尾,因此以 G/C 结尾的使用偏好性在北京丁香质体基因组密码子中是存在的,且色氨酸无密码子偏好性,这一结果与暴马丁香^[41]和辽东丁香^[46]的结果一致。分析北京丁香的密码子使用偏好性对于揭示其遗传信息传递的规律具有重要的研究价值。

基于北京丁香及 36 种植物的质体基因组序列的 NJ 系统发育树研究表明,北京丁香与丁香属植物暴马丁香(*S. reticulata* subsp. *amurensis*)亲缘关系较近,这是因为二者均为短花管组,都属于网脉丁香(*S. reticulata*)的变种,具有及其相似的形态特征,并且二者的花部性状都容易变异,主要靠蒴果

的形状进行区分,同时二者含有相同的染色体数目^[47]。而暴马丁香和北京丁香都与女贞属植物的亲缘关系较近,与其它丁香属植物的关系较远。且所有的丁香属植物与女贞属植物以 100%支持率聚为一支,这是由于两属植物在形态上存在高度的相

似性^[41]。基于质体基因组的系统发育分析,本研究进一步证实了暴马丁香和北京丁香与女贞属植物的密切亲缘关系,这一结论与之前对暴马丁香的质体基因组研究结果高度一致^[41],为理解这两类植物的进化关系提供了有力证据。

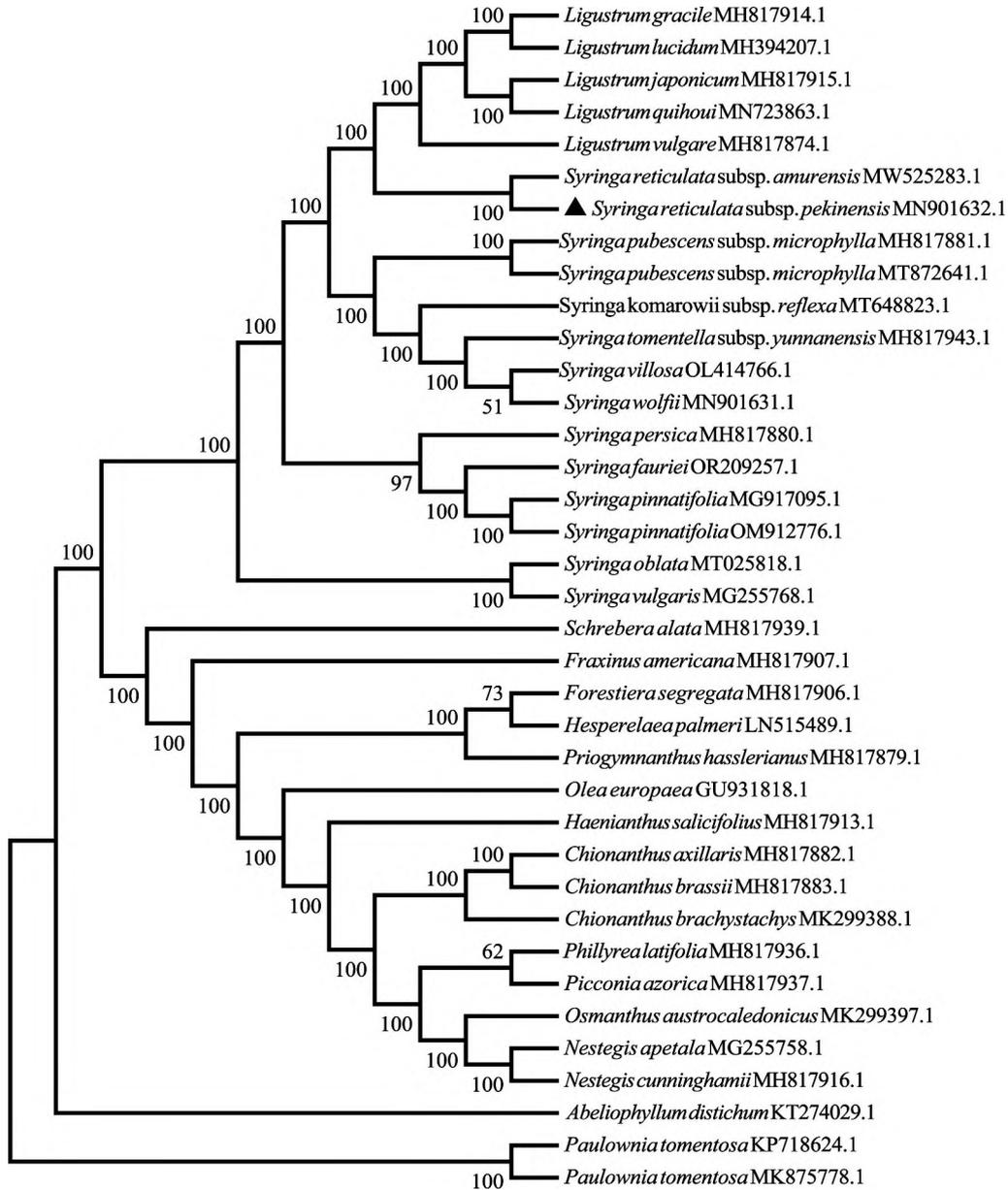


图 5 基于 35 种木犀科植物和 2 种泡桐科植物完整质体基因组的 NJ 系统发育树

Fig.5 NJ phylogenetic tree based on the complete plastid genomes of 35 species of Oleaceae and 2 species of *Paulownia*

4 结论

北京丁香质体基因组全长 155 819 bp,为典型的四分体结构,碱基 GC 含量 38.06%,共注释了 131 个基因,包括 36 个 tRNA 基因,8 个 rRNA

基因和 87 个蛋白质编码基因;在该质体基因组的密码子组成中存在以 G 和 C 结尾的使用偏好性;在北京丁香质体基因组中仅发现正向重复序列和回文重复序列;NJ 系统发育树表明北京丁香与丁香属植物暴马丁香亲缘关系较近,这与传统分类的处置一致,二者都是网脉丁香的亚种;丁香属植

物与女贞属植物亲缘关系较近,女贞属为单系类群,而丁香属为并系类群。

5 致谢

在论文撰写过程中,得到了青海民族大学王久利老师的热心帮助,感谢王老师在绘图及数据分析时给予的指导。

参考文献:

- [1] 中国科学院中国植物志委员会.中国植物志:第六十一卷[M].北京:科学出版社,1992.
- [2] 丁泽英.丁香花在园林中的功能与应用[J].北京农业,2014(21):116.
- [3] 肖战峰.北京丁香种子育苗技术[J].农业科技与信息,2018(4):97.
- [4] 刘云华.谈丁香的药用价值[J].黑龙江生态工程职业学院学报,2011,24(4):41.
- [5] 曲颖,高原,高星宇,等.丁香精油的抑菌作用及其研究进展[J].辽宁化工,2020,49(9):1121-1123.
- [6] 许尚昕,李戴璐,于秋莹,等.北京丁香四倍体诱导及其鉴定[J/OL].分子植物育种,1-10(2023-08-01)[2024-03-20].<http://kns.cnki.net/kcms/detail/46.1068.s.20230817.1340.016.html>.
- [7] 陈长春,张新彩,李万波,等.秦岭11种丁香属植物抗菌作用的实验研究[J].陕西林业科技,1995(3):11-12.
- [8] 宁明世.基于文献的丁香属植物研究进展[J].内蒙古林业科技,2020,46(1):37-42.
- [9] 司兴元,王力刚,王景龙.丁香的繁育及其应用价值[J].防护林科技,2010(3):80-81.
- [10] 杜新宇,卢金梅,李德铎.石松类和蕨类植物物质体基因组结构演化研究进展[J].生物多样性,2019,27(11):1172-1183.
- [11] 李小娟,秦丽莹,韩龙,等.草甸退化指示植物黄帚橐吾的质体基因组特征[J].青海科技,2022,29(3):64-73.
- [12] 温莹莹.兰科天麻族质体基因组进化生物学研究[D].南昌:南昌大学,2023.
- [13] 赵月梅,杨振艳,赵永平,等.木犀科植物叶绿体基因组结构特征和系统发育关系[J].植物学报,2019,54(4):441-454.
- [14] 杨雨青,谭娟,汪芳,等.茶树叶绿体基因组的研究与应用进展[J].生物技术通报,2024,40(2):20-30.
- [15] 樊守金,郭秀秀.植物叶绿体基因组研究及应用进展[J].山东师范大学学报(自然科学版),2022,37(1):22-31.
- [16] BOBLENZ K, NOTHNAGEL T, METZLAFF M. Paternal inheritance of plastids in the genus *Daucus*[J]. *Molecular and General Genetics*, 1990, 220(3):489-491.
- [17] 胡赞民,胡适宜,张金忠.牵牛属质体DNA的父系遗传[J].植物学报,1996,38(4):2-5.
- [18] 陈超南,陆嘉惠,李学禹,等.甘草属种间杂交种叶绿体DNA父系遗传的发现及分析[J].广西植物,2017,37(2):162-168.
- [19] NEALE D B, MARSHALL K A, SEDEROFF R R. Chloroplast and mitochondrial DNA are paternally inherited in *Sequoia sempervirens* D. Don Endl[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1990, 86(23):9347-9349.
- [20] 李泳潭,张军,黄亚丽,等.杜梨叶绿体基因组分析[J].园艺学报,2020,47(6):1021-1032.
- [21] 吴民富,李莎,高柔敏,等.猫尾草叶绿体基因组结构及其解析[J].福建农业学报,2023,38(9):1094-1102.
- [22] 尚明越,王嘉乐,周莹,等.濒危紫皮石斛叶绿体基因组结构及系统发育分析[J].中草药,2023,54(19):6424-6433.
- [23] CHEN S F, ZHOU Y Q, CHEN Y R, et al. Fastp: an ultra-fast all-in-one FASTQ preprocessor[J]. *Bioinformatics*, 2018, 34(17):i884-i890.
- [24] MICHAEL T, PASCAL L, TOMMASO P, et al. GeSeq-versatile and accurate annotation of organelle genomes.[J]. *Nucleic Acids Research*, 2017, 45(W1):W6-W11.
- [25] 胥富强,王海洋,师春娟,等.9种狼尾草属植物叶绿体基因组特征分析[J].西部林业科学,2024,53(1):88-98.
- [26] 原晓龙,李云琴,王毅,等.西南桦叶绿体基因组密码子偏好性分析[J].基因组学与应用生物学,2020,39(12):5758-5764.
- [27] 郭秀丽,王玉,杨路成,等.茶树CBF1基因密码子使用特性分析[J].遗传,2012,34(12):1614-1623.
- [28] 包国媛,李文辛,杨鑫光,等.海甜菜线粒体和叶绿体基因组密码子使用偏好性分析[J].江苏农业学报,2023,39(9):1804-1817.
- [29] 杜明川,王伟,鲍海娟,等.葫芦巴叶绿体基因组密码子偏好性分析[J].草地学报,2024,32(2):409-418.
- [30] 刘慧,王梦醒,岳文杰,等.糜子叶绿体基因组密码子使用偏性的分析[J].植物科学学报,2017,35(3):362-371.
- [31] 杨琳,瞿静涛,李晚忱,等.两种金线莲全长转录本的构建及密码子使用偏好性分析[J].分子植物育种,2020,18(12):3916-3922.
- [32] 武立伟,崔英贤,聂丽萍,等.细茎石斛叶绿体全基

- 因组序列特征及系统发育分析[J].药学学报,2020,55(5):1056-1066.
- [33] 梁璇,李鹏,许冬梅,等.紫苏叶绿体全基因组特征分析[J].山西农业科学,2021,49(3):265-272.
- [34] 杜耀华,王正志.基因组中的重复DNA序列[J].生物技术,2003(6):51-53.
- [35] 张瑞,张天留,宋美华,等.重复序列在牛亚科中的研究进展[J].中国畜牧杂志,2021,57(4):33-38.
- [36] 段义忠,张凯.沙冬青属植物叶绿体基因组对比和系统发育分析[J].西北植物学报,2020,40(8):1323-1332.
- [37] STEFAN K, CHRIS S. REPuter: fast computation of maximal repeats in complete genomes[J].Bioinformatics,1999,15(5):426-427.
- [38] 李思巧,韦伊,刘洪好,等.花椒 cpSSR 标记开发及在种间、种内的通用性分析[J].浙江农林大学学报,2019,36(6):1241-1246.
- [39] PROVAN J, POWELL W, HOLLINGSWORTH P M. Chloroplast microsatellites: new tools for studies in plant ecology and evolution[J].Trends in Ecology & Evolution,2001,16(3):142-147.
- [40] BEIER S, THIEL T, MÜNCH T, et al. MISA - web: a web server for microsatellite prediction. [J].Bioinformatics (Oxford, England),2017,33(16):2583-2585.
- [41] 王久利,郑旭,邓因子.暴马丁香的叶绿体基因组特征分析[J].阜阳师范大学学报(自然科学版),2022,39(1):55-64.
- [42] LIU J, XIONG P P, WANG W P, et al.Characterization of the complete chloroplast genome of *Syringa wolfii*[J].Mitochondrial DNA Part B,2020,5(3):2051-2053.
- [43] CHENG Y Q, JIANG Z M, CAI J. Characteristic and phylogenetic analyses of chloroplast genome for *Syringa komarowii* CK Schneid. (Oleaceae) from Huoditang, China, an important horticultural plant [J].Mitochondrial DNA Part B,2021,6(4):1521-1522.
- [44] ZHAO M D, ZHANG Y M, XIN Z Y, et al.The complete chloroplast genome of *Syringa oblata* (Oleaceae)[J].Mitochondrial DNA Part B,2020,5(3):2278-2279.
- [45] 张靖雯.濒危植物羽叶丁香叶绿体全基因组及系统发育研究[D].杨凌:西北农林科技大学,2019.
- [46] 芦永昌,张鑫,张璐燕,等.辽东丁香完整叶绿体基因组的结构与特征[J].植物研究,2023,43(1):120-130.
- [47] 陈进勇.丁香属(*Syringa* L.)的分类学修订[D].北京:中国科学院研究生院(植物研究所),2006.

Characterization of Chloroplast Genome of *Syringa reticulata* subsp. *pekinensis*

BAO Haijuan^{1,2}, XU Chunliu¹, WANG Yanyan¹, GU Xiaolong¹, LIU Jing^{1,2}

(1. a.College of Ecological Environment and Resources; b.Qinghai Provincial Key Laboratory of Highvalue Utilization of Characteristic Economic Plants; c.Kairui Graduate Workstation of Medicinal Plant Resources, Qinghai Minzu University, Xining 810007, China;

2.Northwest Key Laboratory of Cultivated Land Conservation and Marginal Land Improvement, Ministry of Agriculture and Rural Affairs, Qinghai Bensheng Grass Industry Co., Ltd., Delingha 817000, China)

Abstract: In order to analyze the chloroplast genome characteristics of *Syringa pekinensis*, this study analyzed the chloroplast genome structure and gene composition characteristics, and combined with 35 other Oleaceae species and 2 *Paulownia* species for phylogenetic analysis. The results are as follows. (1) The plastid genome of *S. pekinensis* encodes 131 genes, including 36 tRNA genes, 8 rRNA genes and 87 mRNA genes, accounting for 27.48%, 6.11% and 66.41% of all annotated genes, respectively. (2) In the codon composition of the chloroplast genome, there were 30 codons with RSCU value >1 and 31 codons with RSCU value <1, and the codons had the preference of ending with G and C. (3) A total of 178 forward repeats and 177 palindromic repeats were identified in the plastid genome of *S. pekinensis*. (4) The NJ phylogenetic tree showed that the genetic relationship between *S. pekinensis* and *S. reticulata* was close, and the genetic relationship between *Syringa* and *Ligustrum* was close.

Keywords: *Syringa pekinensis*; Oleaceae; plastid; phylogenetic tree

(责任编辑 李维卫)